MACHINE LEARNING

April 18, 2010

COMS 4771 HOMEWORK #5

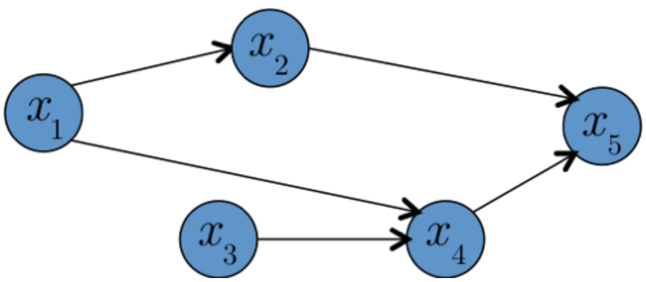
PROF. TONY JEBARA

DUE	MAY 3rd, 2010 BY NOON

Instructions

1. Please follow the same procedure as before for submitting HW5 via Courseworks.

1. (10 points) Directed Graphical Models and Conditional Independence: Consider the directed Bayesian network below. Write out a factorized version of the probability distribution $p(x_1,\ldots,x_5)$ which is implied by the directed graph.



Define the variables as follows:

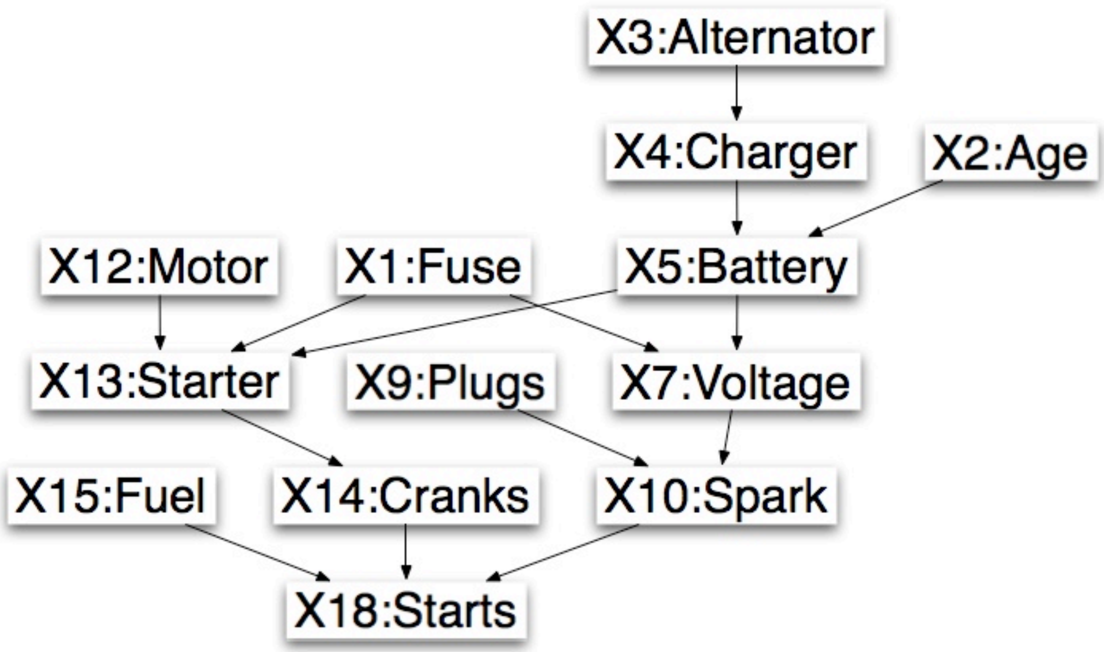
x_1 Student is Intelligent
x_2 Student is Good at Taking Tests
x_3 Student is Hard Working
x_4 Student Understands the Material
x_5 Student Gets Good Grade

Using the Bayes ball algorithm, answer the following questions as True or False.

1. X2 and X4 are independent.
2. X2 and X4 are conditionally independent given X1, X3, and X5.
3. X2 and X4 are conditionally independent given X1 and X3.
4. X5 and X3 are conditionally independent given X4.
5. X5 and X3 are conditionally independent given X1, X2, and X4.
6. X1 and X3 are conditionally independent given X5.
7. X1 and X3 are conditionally independent given X2.
8. X2 and X3 are independent.
9. X2 and X3 are conditionally independent given X5.
10. X2 and X3 are conditionally independent given X5 and X4.

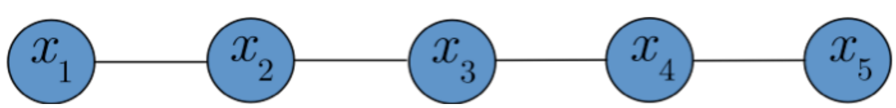
2. (10 points) Junction Tree

You are given the following directed graphical model or Bayesian network for determining if a car will start or not. Assume all variables in the network are binary. Also, assume there are no observed nodes or evidence. Construct the corresponding junction tree by moralizing, triangulating and then building a junction tree that satisfies the running intersection property. BELOW IS A SIMPLIFIED NETWORK.



3. (30 points) Junction Tree Algorithm on a Markov Chain

Consider undirected graphical models of the form shown below which is known as a Markov chain. Assume all variables in the model are binary.



The factorized version of the probability distribution $p(x_1,\ldots,x_5)$ implied by the undirected graph is as follows:

$$p(x) = \frac{1}{Z}\psi(x_1,x_2)\psi(x_2,x_3)\psi(x_3,x_4)\psi(x_4,x_5)$$

Write an implementation of the junction tree algorithm that computes pairwise marginals $p(x_i,x_{i+1})$ for such a Markov chain with any number of variables $p(x_1,\ldots,x_n)$ from the clique potential functions.

The input to your junction tree algorithm should be a cell array of n-1 elements each of which is a 2x2 matrix of non-negative values.

```
n=5;
psis = cell(n-1,1);
for i=1:(n-1)
  psis{i}=rand(2,2);
end
```

The output of your junction tree algorithm (JTA) should be an identical data structure. It should contain consistent marginals that sum to unity appropriately and agree pairwise. Since the tree is only a chain, you don't have to implement a recursive algorithm (i.e. the Collect and Distribute steps in the Jordan book). Instead you only need to perform left to right message passing and then right to left message passing by using a for loop or standard iteration. In other words, the JTA should process the cliques for i=1:n-1 and then for i=n-1:-1:1 to do all the necessary messages in the JTA.

Test your algorithm by recovering the marginals for the following values for the potential functions in the graphical model above.

$$\psi(x_1,x_2) = \begin{array}{c} \\ x_1 = 0 \\ x_1 = 1 \end{array} \begin{array}{cc} x_2 = 0 & x_2 = 1 \\ \multicolumn{2}{c}{\begin{bmatrix} 0.1 & 0.7 \\ 0.8 & 0.3 \end{bmatrix}} \end{array}$$

$$\psi(x_2,x_3) = \begin{array}{c} \\ x_2 = 0 \\ x_2 = 1 \end{array} \begin{array}{cc} x_3 = 0 & x_3 = 1 \\ \multicolumn{2}{c}{\begin{bmatrix} 0.5 & 0.1 \\ 0.1 & 0.5 \end{bmatrix}} \end{array}$$

$$\psi(x_3,x_4) = \begin{array}{c} \\ x_3 = 0 \\ x_3 = 1 \end{array} \begin{array}{cc} x_4 = 0 & x_4 = 1 \\ \multicolumn{2}{c}{\begin{bmatrix} 0.1 & 0.5 \\ 0.5 & 0.1 \end{bmatrix}} \end{array}$$

$$\psi(x_4,x_5) = \begin{array}{c} \\ x_4 = 0 \\ x_4 = 1 \end{array} \begin{array}{cc} x_5 = 0 & x_5 = 1 \\ \multicolumn{2}{c}{\begin{bmatrix} 0.9 & 0.3 \\ 0.1 & 0.3 \end{bmatrix}} \end{array}$$